

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.**

Application Serial Number: 09/638,693A  
Source: IFW/6  
Date Processed by STIC: 2/4/05

# ***ENTERED***



IFW16

## RAW SEQUENCE LISTING

DATE: 02/04/2005

PATENT APPLICATION: US/09/638,693A

TIME: 16:30:31

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\02042005\I638693A.raw

5 <110> APPLICANT: Innogenetics N.V.

9 <120> TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy

10 and therapy

13 <130> FILE REFERENCE: 35

C--> 17 <140> CURRENT APPLICATION NUMBER: US/09/638,693A

C--> 17 <141> CURRENT FILING DATE: 2000-08-15

17 <160> NUMBER OF SEQ ID NOS: 284

21 <170> SOFTWARE: PatentIn version 3.1

25 <210> SEQ ID NO: 1

27 <211> LENGTH: 213

29 <212> TYPE: DNA

31 <213> ORGANISM: hepatitis C virus

35 <400> SEQUENCE: 1

36	ctcacggaac ggctttactg cggggggccct atgttcaaca gcaagggggc ccagtgtggt	60
38	tatcgccgct gccgtgccag tggagttctg cctaccagct tcggcaacac aatcacttgc	120
40	tacatcaagg ccacagcggc tgcaagggcc gcaggcctcc ggaacccgga ctttcttgtc	180
42	tgcggagatg atctggtcgt ggtggctgag agt	213

45 <210> SEQ ID NO: 2

47 <211> LENGTH: 71

49 <212> TYPE: PRT

51 <213> ORGANISM: hepatitis C virus

55 <400> SEQUENCE: 2

57	Leu Thr Glu Arg Leu Tyr Cys Gly Gly Pro Met Phe Asn Ser Lys Gly	
58	1 5 10 15	
61	Ala Gln Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Pro Thr	
62	20 25 30	
65	Ser Phe Gly Asn Thr Ile Thr Cys Tyr Ile Lys Ala Thr Ala Ala Ala	
66	35 40 45	
69	Arg Ala Ala Gly Leu Arg Asn Pro Asp Phe Leu Val Cys Gly Asp Asp	
70	50 55 60	
73	Leu Val Val Val Ala Glu Ser	
74	65 70	

77 <210> SEQ ID NO: 3

79 <211> LENGTH: 213

81 <212> TYPE: DNA

83 <213> ORGANISM: hepatitis C virus

87 <400> SEQUENCE: 3

88	ctcacggaac ggctttactg cggggggccct atgttcaaca gcaagggggc ccagtgtggt	60
90	tatcgccgct gccgtgccag tggagttctg cctaccagct tcggcaacac aatcacttgc	120
92	tacatcaagg ccacagcggc tgcaagggcc gcaggcctcc ggaacccgga ctttcttgtc	180
94	tgcggagatg atctggtcgt ggtggctgag agt	213

97 <210> SEQ ID NO: 4

99 <211> LENGTH: 71

## RAW SEQUENCE LISTING

DATE: 02/04/2005

PATENT APPLICATION: US/09/638,693A

TIME: 16:30:31

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\02042005\I638693A.raw

```

101 <212> TYPE: PRT
103 <213> ORGANISM: hepatitis C virus
107 <400> SEQUENCE: 4
109 Leu Thr Glu Arg Leu Tyr Cys Gly Gly Pro Met Phe Asn Ser Lys Gly
110 1 5 10 15
113 Ala Gln Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Pro Thr
114 20 25 30
117 Ser Phe Gly Asn Thr Ile Thr Cys Tyr Ile Lys Ala Thr Ala Ala Ala
118 35 40 45
121 Lys Ala Ala Gly Leu Arg Asn Pro Asp Phe Leu Val Cys Gly Asp Asp
122 50 55 60
125 Leu Val Val Val Ala Glu Ser
126 65 70
129 <210> SEQ ID NO: 5
131 <211> LENGTH: 213
133 <212> TYPE: DNA
135 <213> ORGANISM: hepatitis C virus
139 <400> SEQUENCE: 5
140 ctcacggagc ggctttactg cggggggccct atgtttaaca gcaagggggc ccagtgtggt 60
142 tategccgtt gccgtgccag tggagttctg cctaccagct tcggcaacac aatcacttgt 120
144 tacatcaaag ccacagcggc cgcaaaagcc gcaggcctcc ggagcccgga ctttcttgtc 180
146 tgcggagatg atctggtcgt ggtggctgag agt 213
149 <210> SEQ ID NO: 6
151 <211> LENGTH: 71
153 <212> TYPE: PRT
155 <213> ORGANISM: hepatitis C virus
159 <400> SEQUENCE: 6
161 Leu Thr Glu Arg Leu Tyr Cys Gly Gly Pro Met Phe Asn Ser Lys Gly
162 1 5 10 15
165 Ala Gln Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Pro Thr
166 20 25 30
169 Ser Phe Gly Asn Thr Ile Thr Cys Tyr Ile Lys Ala Thr Ala Ala Ala
170 35 40 45
173 Lys Ala Ala Gly Leu Arg Ser Pro Asp Phe Leu Val Cys Gly Asp Asp
174 50 55 60
177 Leu Val Val Val Ala Glu Ser
178 65 70
181 <210> SEQ ID NO: 7
183 <211> LENGTH: 213
185 <212> TYPE: DNA
187 <213> ORGANISM: hepatitis C virus
191 <400> SEQUENCE: 7
192 ctcacggagc ggctttactg cggggggccct atgtttaaca gcaagggggc ccagtgtggt 60
194 tategccgtt gccgtgccag tggagttctg cctaccagct tcggcaacac aatcacttgt 120
196 tacatcaaag ccacagcggc cgcaaaagcc gcaggcctcc ggagcccgga ctttcttgtc 180
198 tgcggagatg atctggtcgt ggtggctgag agt 213
201 <210> SEQ ID NO: 8
203 <211> LENGTH: 71
205 <212> TYPE: PRT

```

## RAW SEQUENCE LISTING

DATE: 02/04/2005

PATENT APPLICATION: US/09/638,693A

TIME: 16:30:31

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\02042005\I638693A.raw

207 &lt;213&gt; ORGANISM: hepatitis C virus

211 &lt;400&gt; SEQUENCE: 8

213 Leu Thr Glu Arg Leu Tyr Cys Gly Gly Pro Met Phe Asn Ser Lys Gly

214 1 5 10 15

217 Ala Gln Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Pro Thr

218 20 25 30

221 Ser Phe Gly Asn Thr Ile Thr Cys Tyr Ile Lys Ala Thr Ala Ala Ala

222 35 40 45

225 Lys Ala Ala Gly Leu Arg Ser Pro Asp Phe Leu Val Cys Gly Asp Asp

226 50 55 60

229 Leu Val Val Val Ala Glu Ser

230 65 70

233 &lt;210&gt; SEQ ID NO: 9

235 &lt;211&gt; LENGTH: 213

237 &lt;212&gt; TYPE: DNA

239 &lt;213&gt; ORGANISM: hepatitis C virus

243 &lt;400&gt; SEQUENCE: 9

244 ctacaggagc ggctttactg cgggggccct atgttcaaca gcaagggggc ccagtgtggt 60

246 tategccgtt gtcgtgccag tggagttctg cctaccagtt tcggcaacac aatcacttgt 120

248 tacatcaagg ccacagcggc tgcaaaagcc gcaggcctcc ggaacccgga ctttcttggt 180

250 tgcggagatg atttggtcgt ggtggctgag agt 213

253 &lt;210&gt; SEQ ID NO: 10

255 &lt;211&gt; LENGTH: 71

257 &lt;212&gt; TYPE: PRT

259 &lt;213&gt; ORGANISM: hepatitis C virus

263 &lt;400&gt; SEQUENCE: 10

265 Leu Thr Glu Arg Leu Tyr Cys Gly Gly Pro Met Phe Asn Ser Lys Gly

266 1 5 10 15

269 Ala Gln Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Pro Thr

270 20 25 30

273 Ser Phe Gly Asn Thr Ile Thr Cys Tyr Ile Lys Ala Thr Ala Ala Ala

274 35 40 45

277 Lys Ala Ala Gly Leu Arg Asn Pro Asp Phe Leu Val Cys Gly Asp Asp

278 50 55 60

281 Leu Val Val Val Ala Glu Ser

282 65 70

285 &lt;210&gt; SEQ ID NO: 11

287 &lt;211&gt; LENGTH: 213

289 &lt;212&gt; TYPE: DNA

291 &lt;213&gt; ORGANISM: hepatitis C virus

295 &lt;400&gt; SEQUENCE: 11

296 ctacaggagc ggctttactg cgggggccct atgttcaaca gcaagggggc ccagtgtggt 60

298 tategccgtt gtcgtgccag tggagttctg cctaccagtt tcggcaacac aatcacttgt 120

300 tacatcaagg ccacagcggc tgcaaaagcc gcaggcctcc ggaacccgga ctttcttggt 180

302 tgcggagatg atttggtcgt ggtggctgag agt 213

305 &lt;210&gt; SEQ ID NO: 12

307 &lt;211&gt; LENGTH: 71

309 &lt;212&gt; TYPE: PRT

311 &lt;213&gt; ORGANISM: hepatitis C virus

## RAW SEQUENCE LISTING

DATE: 02/04/2005

PATENT APPLICATION: US/09/638,693A

TIME: 16:30:31

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\02042005\I638693A.raw

315 &lt;400&gt; SEQUENCE: 12

317 Leu Thr Glu Arg Leu Tyr Cys Gly Gly Pro Met Phe Asn Ser Lys Gly

318 1 5 10 15

321 Ala Gln Cys Gly Tyr Arg Arg Cys Arg Ala Ser Gly Val Leu Pro Thr

322 20 25 30

325 Ser Phe Gly Asn Thr Ile Thr Cys Tyr Ile Lys Ala Thr Ala Ala Ala

326 35 40 45

329 Lys Ala Ala Gly Leu Arg Asn Pro Asp Phe Leu Val Cys Gly Asp Asp

330 50 55 60

333 Leu Val Val Val Ala Glu Ser

334 65 70

337 &lt;210&gt; SEQ ID NO: 13

339 &lt;211&gt; LENGTH: 541

341 &lt;212&gt; TYPE: DNA

343 &lt;213&gt; ORGANISM: hepatitis C virus

347 &lt;400&gt; SEQUENCE: 13

348 cgctggcgct cctgtaggag gcgtcgcaag agcccttgcg catggcggtga gggcccttga 60

350 agacgggata aatttcgcaa caggggaattt gcccggttgc tccttttcta tcttccttct 120

352 tgctctgttc tcttgcttaa tccatccagc agctagtcta gaggggcgga acacgtctgg 180

354 cctctatgtc cttaccaacg actgttccaa tagcagtatt gtgtatgagg ccgatgacgt 240

356 tattctgcac acacccggct gtgtaccttg tgttcaggac ggtaatacat ctgcgtgctg 300

358 gacccagtg acacctacag tggcagtcag gtacgtcgga gcaaccaccg ctcgatacg 360

360 caggcatgta gacatgttgg tgggcgcggc cacgatgtgc tctgctctct acgtgggtga 420

362 tatgtgtggg gccgtcttcc tcgtgggaca agccttcacg ttcagacctc gtcgccatca 480

364 aacgggtccag acctgtaact gctcactgta cccaggccat ctttcaggac accgaatggc 540

366 t 541

369 &lt;210&gt; SEQ ID NO: 14

371 &lt;211&gt; LENGTH: 180

373 &lt;212&gt; TYPE: PRT

375 &lt;213&gt; ORGANISM: hepatitis C virus

379 &lt;400&gt; SEQUENCE: 14

381 Val Gly Ala Pro Val Gly Gly Val Ala Arg Ala Leu Ala His Gly Val

382 1 5 10 15

385 Arg Ala Leu Glu Asp Gly Ile Asn Phe Ala Thr Gly Asn Leu Pro Gly

386 20 25 30

389 Cys Ser Phe Ser Ile Phe Leu Leu Ala Leu Phe Ser Cys Leu Ile His

390 35 40 45

393 Pro Ala Ala Ser Leu Glu Trp Arg Asn Thr Ser Gly Leu Tyr Val Leu

394 50 55 60

397 Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr Glu Ala Asp Asp Val

398 65 70 75 80

401 Ile Leu His Thr Pro Gly Cys Val Pro Cys Val Gln Asp Gly Asn Thr

402 85 90 95

405 Ser Ala Cys Trp Thr Pro Val Thr Pro Thr Val Ala Val Arg Tyr Val

406 100 105 110

409 Gly Ala Thr Thr Ala Ser Ile Arg Arg His Val Asp Met Leu Val Gly

410 115 120 125

413 Ala Ala Thr Met Cys Ser Ala Leu Tyr Val Gly Asp Met Cys Gly Ala

414 130 135 140

## RAW SEQUENCE LISTING

DATE: 02/04/2005

PATENT APPLICATION: US/09/638,693A

TIME: 16:30:31

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\02042005\I638693A.raw

```

417 Val Phe Leu Val Gly Gln Ala Phe Thr Phe Arg Pro Arg Arg His Gln
418 145          150          155          160
421 Thr Val Gln Thr Cys Asn Cys Ser Leu Tyr Pro Gly His Leu Ser Gly
422          165          170          175
425 His Arg Met Ala
426          180
429 <210> SEQ ID NO: 15
431 <211> LENGTH: 541
433 <212> TYPE: DNA
435 <213> ORGANISM: hepatitis C virus
439 <400> SEQUENCE: 15
440 cgtcggcgct cctgtaggag gcgtcgcaag agcccttgcg catggcgtga gggcccttga      60
442 agacgggata aatttcgcaa caggggaattt gcccggttgc tccttttcta tcttccttcc      120
444 tgctctgttc tcttgcttaa tccatccagc agctagtcta gagtggcgga acacgtctgg      180
446 cctctatgtc cttaccaacg actgttccaa tagcagtatt gtgtatgagg ccgatgacgt      240
448 tattctgcac acacccggct gtgtaccttg tgttcaggac ggtaatacat ctgcgtgctg      300
450 gacccagtg acacctacag tggcagtcag gtacgtcgga gcaaccaccg cttcgatacg      360
452 caggcatgta gacatattgg tgggcgcggc cacaatgtgc tctgctctct acgtgggtga      420
454 tatgtgtggg gccgtcttcc tcgtgggaca agccttcacg ttcagacctc gtcgccatca      480
456 aacggtccag acctgtaact gctcactgta cccaggccat ctttcaggac accgaatggc      540
458 t
461 <210> SEQ ID NO: 16
463 <211> LENGTH: 180
465 <212> TYPE: PRT
467 <213> ORGANISM: hepatitis C virus
471 <400> SEQUENCE: 16
473 Val Gly Ala Pro Val Gly Gly Val Ala Arg Ala Leu Ala His Gly Val
474 1          5          10          15
477 Arg Ala Leu Glu Asp Gly Ile Asn Phe Ala Thr Gly Asn Leu Pro Gly
478          20          25          30
481 Cys Ser Phe Ser Ile Phe Leu Pro Ala Leu Phe Ser Cys Leu Ile His
482          35          40          45
485 Pro Ala Ala Ser Leu Glu Trp Arg Asn Thr Ser Gly Leu Tyr Val Leu
486          50          55          60
489 Thr Asn Asp Cys Ser Asn Ser Ser Ile Val Tyr Glu Ala Asp Asp Val
490 65          70          75          80
493 Ile Leu His Thr Pro Gly Cys Val Pro Cys Val Gln Asp Gly Asn Thr
494          85          90          95
497 Ser Ala Cys Trp Thr Pro Val Thr Pro Thr Val Ala Val Arg Tyr Val
498          100         105         110
501 Gly Ala Thr Thr Ala Ser Ile Arg Arg His Val Asp Ile Leu Val Gly
502          115         120         125
505 Ala Ala Thr Met Cys Ser Ala Leu Tyr Val Gly Asp Met Cys Gly Ala
506          130         135         140
509 Val Phe Leu Val Gly Gln Ala Phe Thr Phe Arg Pro Arg Arg His Gln
510 145          150          155          160
513 Thr Val Gln Thr Cys Asn Cys Ser Leu Tyr Pro Gly His Leu Ser Gly
514          165          170          175
517 His Arg Met Ala

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/638,693A

DATE: 02/04/2005  
TIME: 16:30:32

Input Set : A:\PTO.FG.txt  
Output Set: N:\CRF4\02042005\I638693A.raw

**Please Note:**

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:55; N Pos. 191  
Seq#:56; Xaa Pos. 64  
Seq#:57; N Pos. 185,191,276  
Seq#:58; Xaa Pos. 62,64,92  
Seq#:59; N Pos. 38,76,79,151,232,349  
Seq#:60; Xaa Pos. 13,26,27,51,78,117  
Seq#:61; N Pos. 38,70,174,339  
Seq#:62; Xaa Pos. 13,24,58,113  
Seq#:165; N Pos. 306,313,340  
Seq#:166; Xaa Pos. 102,105,114  
Seq#:179; N Pos. 568,569,570,571,572,573,574,575,576,577,578,579  
Seq#:180; Xaa Pos. 24,190,191,192,193  
Seq#:181; N Pos. 152  
Seq#:182; Xaa Pos. 51  
Seq#:197; N Pos. 787,1011,1444  
Seq#:198; Xaa Pos. 263,337,482  
Seq#:199; N Pos. 684,909,1058,1133  
Seq#:200; Xaa Pos. 228,303,353,378  
Seq#:209; N Pos. 13,23,65,72,277,337  
Seq#:210; Xaa Pos. 8,22,24  
Seq#:211; N Pos. 11,327,333,334,340  
Seq#:212; Xaa Pos. 4,109,111

**Invalid <213> Response:**

Use of "Artificial" only as "<213> Organism" response is incomplete,  
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,124,125,141  
Seq#:142

## VERIFICATION SUMMARY

DATE: 02/04/2005

PATENT APPLICATION: US/09/638,693A

TIME: 16:30:32

Input Set : A:\PTO.FG.txt

Output Set: N:\CRF4\02042005\I638693A.raw

L:17 M:270 C: Current Application Number differs, Replaced Current Application No  
L:17 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:2300 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:55 after pos.:180  
L:2341 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:56 after pos.:48  
L:2408 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:57 after pos.:180  
M:341 Repeated in SeqNo=57  
L:2469 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:58 after pos.:48  
M:341 Repeated in SeqNo=58  
L:2564 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:59 after pos.:0  
M:341 Repeated in SeqNo=59  
L:2639 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:60 after pos.:0  
M:341 Repeated in SeqNo=60  
L:2722 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:61 after pos.:0  
M:341 Repeated in SeqNo=61  
L:2787 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:62 after pos.:0  
M:341 Repeated in SeqNo=62  
L:5642 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:165 after pos.:300  
L:5717 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:166 after pos.:96  
M:341 Repeated in SeqNo=166  
L:6340 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:179 after pos.:540  
L:6379 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:180 after pos.:16  
M:341 Repeated in SeqNo=180  
L:6452 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:181 after pos.:120  
L:6503 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:182 after pos.:48  
L:7234 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:197 after pos.:780  
M:341 Repeated in SeqNo=197  
L:7365 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:198 after pos.:256  
M:341 Repeated in SeqNo=198  
L:7498 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:199 after pos.:660  
M:341 Repeated in SeqNo=199  
L:7635 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:200 after pos.:224  
M:341 Repeated in SeqNo=200  
L:8046 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:209 after pos.:0  
M:341 Repeated in SeqNo=209  
L:8101 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:210 after pos.:0  
M:341 Repeated in SeqNo=210  
L:8184 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:211 after pos.:0  
M:341 Repeated in SeqNo=211  
L:8239 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:212 after pos.:0  
M:341 Repeated in SeqNo=212